On the methodology of network analysis in sociology

Balzer, Wolfgang; Flap, Hendrik Derk

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

ON THE METHODOLOGY OF NETWORK ANALYSIS IN SOCIOLOGY*

by

W. BALZER
Seminar for Philosophy, Logic and the Philosophy of Science,
University of Munich, Ludwigstr. 31, Munich 22, D-8000, W. Germany

and

H.D. FLAP

Dept. of Sociology, University of Utrecht, Heidelberglaan 2,
P.O.Box 80.108, 3508 TC Utrecht, Netherlands

Introduction

In this paper we will look at sociological network analysis from the point of view of recent developments in the philosophy of science. Although at the moment network analysis represents an optimal way of doing empirical work in sociology, we will argue that network analysis still is in a premature state. We will corroborate this by analyzing its formalism, cleavages, positions and organizations. It will become clear that achievements so far do not cover an adequate, unified theory. We do not intend to criticize mainly the approach. On the contrary: we are looking for ways that may lead towards a comprehensive theory of social networks. In doing so we will draw attention to the role played by theoretical terms and underlying theories.

What is Network Analysis?

During the last two decades the body of writings on social networks has been considerable. R.S. Burt (1980) gives the reader in his 'Models

* The joint work at this paper was made possible by fellowships of the Netherland Institute for Advanced Studies (NIAS).

of Network Structure' a good survey with many references. Various features are common to all:

1. There is a common use of graphs, matrices, networks etc. as pictorial models representing real, sociologically relevant features of the world. The picture of a net e.g. consists of knots (representing persons or groups) which are connected by lines (representing relations among persons or groups). Various local or global features of such formal networks in concrete situations correspond to clearly identifiable features of human relations (given fixed interpretations of the knots and lines in the net). If, for example, the lines denote a flow of information from one person to another, and a net is such that all lines from one person to any other are passing through one distinguished knot, then the person denoted by that knot in the corresponding concrete situation has complete control over any flow of information in the set of persons considered.

2. Most authors assume that the number of relations of a given type in which a person or a group is involved is a grade of importance of that person or group with respect to the given type of relation. If a person as a friend (or other type of relation) has relations to thirty other persons while another individual has such relations with only three other individuals then the former is more important in the system than the latter, and that special type of relation is more important for the former as well. There are many investigations during which numbers of relations are counted. But counting is associated with measurement, and thus with giving different 'weight' or 'importance' to different persons.

3. Most writings show an operationalist or behaviorist metaphysical background. The very idea of interpreting graphs by means of persons and their relations seems to be free of general, theoretical assumptions. Persons can be directly observed; their relations are open to direct inspection. The investigation can start with the analysis of overt behavior.

4. A fourth feature, relevant to our discussion is also an item which may be overlooked easily: The method of graphs is applied to different situations or types of phenomena. Assumptions assumed to be valid in those special contexts differ widely. There is no fixed set of assumptions that is satisfied in all applications of the method.

5. Last but not least there prevails among practitioners of network methods a conviction that empirical conclusions (predictions) are the outcome of relations among *different* networks. However, we cannot predict individual behavior or individual relations knowing only that a certain individual is a member of a certain network. Predictions regarding future events are as a rule based on the knowledge of how different networks to which this individual belongs can be brought into line, and how they are interrelated.

These features appear as a *paradigm* in the sense of Thomas Kuhn. Sociological network research can be clearly identified and delineated. It is not necessary, however, to engage in mapping the above features on conditions as described by Kuhn.[1] His characterizations are much too weak and too ambivalent, and of minor importance to our present considerations.

On the other hand it is quite clear that the afore-said features are not sufficient to meet with the demands of a *theory*. It is common to all theory-concepts which are discussed in the literature that a theory consists – among other things – of a certain vocabulary (basic notions), and some axioms and assumptions (sentences formulated in the vocabulary).[2] None of the above mentioned features includes proper axioms, perhaps with the exception of the axioms characterizing mathematical graphs.

We might try to introduce a system of axioms or a class of models for network theory. However, from the point of view of recent developments in the philosophy of science this would not yield an adequate picture. For network theory is a comprehensive theory, and the picture of comprehensive theories which has emerged recently is that of a *theory-net*[3], consisting of a set of 'little' theories or laws (as the knots of the net) together with a specialization relation between those laws (representing the lines), expressing that one law is a specialization of another one. In such a net there is one distinguished knot which repre-

1. Especially in his (Kuhn, 1970), appendix.
2. In set-theoretic or 'structuralist' accounts the vocabulary is represented by the 'type' of the models and the axioms are represented by a corresponding class of models. Compare Balzer & Heidelberger, 1983 for the different theory-concepts available.
3. Compare Balzer & Sneed, 1977/78.

sents the basic laws or assumptions of the theory, that is, those assumptions which hold in *all* applications of the theory. Usually, such basic laws are rather trivial due to the fact that they are valid in so many different applications of very different nature. Examples are Newton's second law as the basic law of classical mechanics, and Maxwell's equations as the basic law of electrodynamics. The basic laws constitute a kind of general frame which guides the search for special laws. Special laws are laws which are valid only in a subset of the set of all applications of the theory, they only apply to special systems. Hooke's law e.g. does not hold for very great values of stress: It does not give a complete description of the elastic behavior of a substance. Different special laws are 'bound together' by being specializations of one common basic law. This assures that the same basic notions are applied in all cases, and therefore that the different applications, in fact, belong to one theory. We believe that this picture of theory-nets is also relevant for network theory. That is, a satisfactory account of network theory is likely to exemplify some basic laws for social networks holding in general as well as further special laws being characteristic for special applications (focussing in special features like cleavages, positions, etc.).

Usually, the establishment of a paradigm leads to the introduction of a general theory. At least this is what we learn from the history of science. If this is true then in the case of network analysis, too, we should expect the development and introduction of a comprehensive theory having all or most of the present, various applications as special cases. Such a theory, comparable to, say Newtonian mechanics, in classical physics, is what we would call a mature theory or a mature state of network analysis. It is in this sense that we claimed network analysis still to be in a premature state. This simply means that there does not yet exist a set of basic axioms for network theories in general such that the various different applications can be obtained by specializations of the general axioms.

We will now turn to more special issues which, we think, are of relevance for the construction of such a comprehensive theory.

Enlarging the Formalism

The formalism inherent in the picture of mathematical graphs or

networks imposes some restrictions which cannot be expected to be met in all examples.

A first restriction is that only *binary* relations are used. The picture of a graph in which each line always connects two points with each other, and also the mathematical definition of a graph as a set of ordered pairs (eventually with additional properties) can only be interpreted as representing binary relations. 'X is a friend of Y' and 'X and Y work together' are typical examples. But it is clear that relations of different types occur in reality and are important as well. Consider for instance the relation 'X is a child of Y and Z' or 'X is a business partner of Y and Z'. These are ternary relations which certainly might be relevant in certain contexts, but which cannot be reduced to binary ones.[4] A more complicated example involving economical concepts would be 'X at t exchanges quantity A of commodity i against quantity B of commodity j'. This example also would necessitate a change of the interpretation of the knots in the net: knots would represent persons, equipped with certain quantities of kinds of commodities at certain times.

What can be seen from these examples is the following: There are other relations, relations with more than two argument-places, which are important and do not seem to be reduceable to binary ones. The application of network analysis in cases where such relations are present cannot be adequately depicted by means of graphs in the present interpretation. In the present interpretation the picture of knots and lines between them indicates that all relations among the knots can be analyzed in terms of binary relations. But this does not fit with sociological reality. There are relations which have to be represented by more than two knots which, together with lines among them, form an inseparable unit.

This is not to say that we are in favor of giving up the graphical representation of social networks as graphs. We simply want to point out that *in general* this method is too restrictive. So, in order to formulate a general theory of social networks, the concept of a mathematical

4. This is not meant as a non-possibility claim in any strict logical sense. Often, logics reveals possibilities which were not discovered by common sense, and it usually is very difficult to substantiate and to 'prove' non-possibility results of the kind mentioned.

graph cannot be a constitutive element. As a rule one is obliged to use the abstract concept of a *relational structure*, as a structure consisting of one or several set(s) of 'objects' together with various relations (or functions) of various argument-places. Relational structures can be precisely understood as structures for many-sorted, higher order languages in the sense of formal logics.[5]

A second restriction which we find unnecessary is the restriction to types of relations such that the number of relations of a given type in which a person is involved indicates some weight. Of course, in special cases relations of the same type may add up to something very important. But mere counting of relations seems to be a rather blind method. As a rule, quantity is no substitute for quality. We conclude that the matrix formalism which is frequently used in order to handle quantitative aspects as given by numbers of relations represents only a special device working nicely in some special cases.

Finally, what is still missing are systematic investigations about relations *across* different nets. Let us compare the situation with lattice theory. In a lattice there are (in general) various possibilities for sub-lattices, and much theoretical work has been devoted to the investigation of relations among sub-lattices. Similar investigations are necessary in the field of network analysis. There certainly is a vast area of different types or ways of how networks can overlap and can be combined. Certainly there will be standard configurations which are especially relevant to sociological applications, for example with respect to the kind of qualitative predictions derived from the way in which different networks overlap for one person.

Theoretical Terms

Before turning to cleavages, positions and organizations we draw your attention to that body of generalizations and principles known as 'theoretical terms' as developed in association with practice, because recently this concept underwent a metamorphosis.[6]

5. Compare Bourbaki, 1968, pp. 262 for precise definitions. There, relational structures are called "structures of species S on the principal base sets $E_1, \ldots, E_n$ with $A_1, \ldots, A_m$ as auxiliary base sets".
6. Compare Balzer, 1983, Sec. V, for a precise presentation of the new definition of theoretical terms.

Traditionally, 'theoretical terms' were regarded as terms which are not 'observable'. They had a derived status, their use was justified only insofar as they could be reduced in meaning to, or defined by, 'non-theoretical' terms. The well known difficulties of this reductionist program which started in logical empiricism, led to a directing of the attention toward single, isolated theories. In a given theory we can draw a similar distinction. Some terms of the theory may be treated as observable with respect to that theory, that is, their meaning can be determined either 'directly' or by means of other theories, but without the help of the theory in question. On the other hand, one might call a term *'theoretical'* in a given theory if it cannot be determined without using the theory itself. It turns out that this idea can be made precise by requiring that a theoretical term is one which is 'weakly definable' in a theory, that is, if at least in some models of the theory the term's denotation is uniquely determined by means of the other terms' denotations. More precisely, the picture is this. In a model of the theory each term has a denotation or interpretation. Let us call term t's denotation in model x *the relation of* x *associated with* t. Then t is weakly definable in theory T if there are models x of T in which the relation of x associated with t is uniquely determined by the other relations of x (i.e. the relations of x associated with T's other terms), and by x being a model of T.

'Weak definability' in this sense has to be contrasted with the usual concept of explicit definability. The central difference is that an explicit definition guarantees the term's eliminability in all contexts in which the theory applies, i.e. in all models of the theory. This is not so for 'weak definability'. A term may be 'weakly definable', but still, in some applications of the theory, it may not be eliminable. For there may be applications which are represented by models in which the denotation of the term is not uniquely determined.

It may be objected that this notion of 'weak definability' is in fact so weak that all terms will be weakly definable, even the non-theoretical ones. In order to prohibit this triviality it is necessary to use a notion of definition slightly stronger than that of 'weak definability'. What we have in mind is a concept which might be called 'definability in the range of invariance of the theory'. This notion is technically a bit involved and

need not be presented here.[7] The only important point for us is that there is a completely precise distinction (which can be fully formalized), and that this distinction in case of 'real-life' theories turns out to be identical with what commonly is accepted as the distinction between theoretical and non-theoretical terms. The basic property of theoretical terms which follows from this definition is that they can be determined by means of the theory under *suitable* conditions, but usually *not* in general, i.e. not in all applications. Non-theoretical terms – in contrast – cannot be determined at all by means of the theory under consideration. They have to be presupposed as 'given', and they acquire their meaning in other theories, different from the one under consideration. We note that in a certain way those other theories which are necessary in order to give some meaning to theory T's 'non-theoretical' terms may be said to be theories *underlying* T. By applying the distinction between theoretical and non-theoretical terms in each of T's underlying theories we obtain further theories underlying the former, namely those in which the non-theoretical terms of the former theories acquire their meaning. And so on. In this way the new definition can be used to trace back and to reconstruct the web of theories underlying one particular, given theory T.

Cleavages

A cleavage may be defined as a partition of a population into two or more subsets. Sometimes the subsets are required to be disjoined and exhaustive, but these requirements do not seem to be regarded as essential by all authors.

Often, cleavages are operationally defined. That is, there are given criteria which, if applied, yield a decision about whether or not an individual belongs to one of the subsets of the partition. For example, if the cleavage consists of two religious groups then a criterion for some individual to be member of one group might simply consist in its being inscribed in one of the commons. The operation by means of which we

7. See Balzer, 1983, Sec. V. What is called a 'term definable in the range of invariance of the theory' here is called a 'T-theoretical term' in the paper mentioned.

decide on membership in this case is to check the various books of the different commons.

If treated in this operationalist way there is no intrinsic connection between the relations of a social network and cleavages. Cleavages are then regarded as being detectable independently of social relations among individuals. But there is also the opposite view according to which cleavages are and have to be explicitly defined in terms of the relations of a social network. The idea is that among the individuals belonging to one common subset of the cleavage there are certain characteristic relations which do not obtain in other subsets of the cleavage or across the subsets. Consideration of examples, however, shows that such definitions usually are not feasible. How should we define, for instance, a person's being member of a religious community merely in terms of that person's individual relations? Or, to take another example, how should we define a person's being member of the 'working class' merely in terms of that person's relations to other individuals. We certainly will find relations *indicating* such a membership, for example frequent contact with the respective priest, or no contact with the owner of the factory, but such indications can never serve as general definitions. In our view both these extreme cases, namely treating cleavages as operationally defined or as explicitly definable in terms of relations, are inadequate. The right way to look at the concept of a cleavage is to regard this concept as a theoretical term. That is, the concept of a cleavage is not in general (i.e. for all possible applications) definable in terms of relations, but it is weakly definable in some special applications. There exist special applications of the network approach in which it is possible under further special assumptions in fact to define cleavages in terms of the relations under consideration. But such a definition will not be adequate in other applications of the network approach.

A theoretical term in the sense introduced in the previous section can occur only in a proper theory. So the term 'cleavage' can – properly speaking – be theoretical only in some network theory which, as indicated earlier, still does not yet exist. Nevertheless, the idea of cleavages being theoretical may serve as a guide or constraint in setting up such a theory.

Positions

What has been said about cleavages applies mutatis mutandis to positions, too. Positions are distinguished places in society connected with special roles, obligations, rights etc. Again, we have two extreme possibilities. On the one hand we may regard a position as operationally defined by means of certain criteria. For instance, we can say that the position of a head of department is defined by certain contracts and regulations of law. In order to find out whether some individual keeps such a position we have to apply the operation consisting of checking whether all relevant contracts and regulations in fact apply to this person.

On the other hand there is the idea of a general definition of *positions* in terms of network relations. Positions are then defined abstractly as equivalence classes of networks with a special internal structure. For instance, we might consider a network consisting of four individuals and lines among individual 1 and i for i = 2, 3, 4. In such a concrete net we can say that individual number 1 occupies a distinguished concrete position. This position will be the same in all nets isomorph to the original one. So if we consider the equivalence class of all nets isomorph to the original one we may say that this equivalence class expresses the position under discussion in an 'abstract' way.[8]

Again, we think that both views, namely of positions as being operationally defined or as being explicitly defined are inadequate. Positions – like cleavages – are theoretical with respect to network theory. That is, in special applications of the theory (which still has to be properly formulated) it will be possible to determine or to define in a weak sense certain special positions, but this determination does not work in all applications of the theory.

Organizations

Organizations are difficult to define, and we will not try to elaborate on one of the common characterizations. We think that all of them are

8. Two networks are isomorph if there is a unique correspondence between the set of knots of the one network and the set of knots of the other network. That is, if two knots in the one network are connected by some line then their images in the other network are also connected by some line.

deficient in various respects. Instead, it seems more adequate to start with clearcut examples. Typical examples are firms in the productive sector, units of bureaucracy (like ministeries, departments of administration), churches, unions, police divisions, courts, as well as small enterprises run by a family.

Here again, we have the possibilities of either defining organizations operationally in a way independent of the relations among the individuals involved or by explicitly defining organizations in terms of social relations.

The first alternative will essentially work with criteria using distinctions from economics, politics, and law. For instance, a firm might be characterized as an abstract actor which pursues profit by means of producing commodities in a certain setting governed by a special political frame and by special laws. Especially in the sphere of law organizations are often treated as person of law so that many authors use anthropomorphic terms like 'corporate actor' instead of the term 'organization'. A moment's reflection shows that operational characterizations like the one just mentioned hardly provide more than a first, rough orientation. All the terms involved in the characterization are highly problematic and unclear. The problem of what is an actor straightforwardly leads into the theory of action with all its dependencies on psychology and philosophy of language. The concept of profit leads into problems of economics, not to speak about the political frame and the law. Of course, one might envisage a theory of organizations as forming the top of a little hierarchy of theories (or sciences) consisting of psychology, economics, philosophy, political science, and law, to mention only the most important parts. But the practical implications of such a picture are severe. A proper theory of organizations would essentially rely on all its underlying theories, as listed above, and therefore always would rely in its quality and degree of precision on the quality and precision of its underlying theories.

This situation yields an argument for the second possibility which consists of explicitly defining organizations in terms of the relations of the individuals involved. Intuitively, this idea is simple and attractive. After all, real life organizations essentially consist of human beings which behave according to certain, rather strict rules, among which various

essential relations hold. This picture avoids also to rely too much on other, underlying theories. Even in organizations, social relations can be directly observed, and thus the observational basis for a theory of organizations based on human relations need not rely on other social sciences. However, this approach does not exist. At the moment, some first, small steps are made in that direction. This may have several reasons. First, the number of relevant relations in an organization usually is rather large, and the relations themselves cover quite different aspects of human life. For instance, the spectrum of relations may range from economic relations of exchange (of commodities against money, or norms against money, or rights against information and so on) on the one hand to strictly delineated relations of power (like the power to give orders of a certain form to certain persons, to be obliged to deliver certain informations, to have the right to apply for certain things and so on).

Second, these various kinds of relations may have to combine in certain specific ways in order to yield an organization. Even the personal enemies of some bureaucrat may use their right to apply to certain things via this person. The political opinions of somebody working in a firm may be tolerated as long as the person contributes efficiently to the firm's goals. And so on. It is clear that this is a big field for research, and in the frame of the present paper we cannot do any general theoretical proposals. However, we want to suggest that the concept of an organization, too, has or should have the status of a theoretical term in network theory. In certain simple organizations, like the family enterprise, one would expect the possibility of a weak definition of such an organization in terms of the relations of the persons involved. But such a weak definition certainly will not do in general for all applications of network theory. The larger the organization the more difficult it will be to achieve a complete characterization in terms of social relations.

So the terms 'cleavage', 'position' and 'organization' all are and should be theoretical terms of network theory. A network theory with such theoretical terms seems to represent a rather substantial extension of the present state of affairs. Network analysts may think of this as a too grandiose program. But against this attitude we want to point out that the three concepts under discussion are central for any of the more

interesting theoretical propositions in sociology. Cleavages as macro-distinctions exist and may be easily detected. By referring to such distinctions sociology is freed from the pressure to start from the individualistic basis with all its difficulties and irregularities. Positions are the right means to give social networks an interesting and non-trivial structure which on the one hand is accessible to an understanding from the point of view of the individual, and on the other hand is abstract enough to allow for general patterns in various different applications. Organizations, finally, perhaps represent the most characteristic feature of present societies.[9] A sociological theory neglecting organizations will not be able to say anything interestingly which is typical for present day societies.

Summary

It is argued that the network approach in sociology – which forms a paradigm in the sense of T. Kuhn – does not yet constitute a mature theory in the sense of modern philosophy of science. The present achievements of the network approach are contrasted with distinct features of mature theories, like e.g. their explicit form of so-called 'theory-nets', or the distinction between theoretical and non-theoretical terms. An outline of a new distinction between theoretical and non-theoretical terms in a theory is given. The methodological status of the terms 'cleavage', 'position' and 'organization' is discussed, and it is argued that these three terms should have the status of theoretical terms in a proper theory of social networks.

9. See for instance Coleman, 1974.

W. Balzer was born in Darmstadt, Germany, in 1947. He got degrees in mathematics and philosophy of science. Since 1976 he teaches philosophy of science. In 1984 he became professor of logics and philosophy of science at Munich university. He published various articles and books, e.g. Empirische Theorien: Modelle, Strukturen, Beispiele.

Hendrik Derk Flap was born in Lutjegast, The Netherlands in 1950. He graduated in sociology at the State University of Groningen in 1974. Presently he is working at the State University of Utrecht, teaching courses in theoretical sociology and the philosophy of science. His main research interest concerns the theoretical underpinning of social network research. He wrote in Kölner Zeitschrift für Soziologie und Sozialpsychologie, 1981.

REFERENCES

Balzer, W., 'Theory and Measurement', *Erkenntnis, 19,* 1983, pp. 3-25.
Balzer, W. & M. Heidelberger (Eds.), *Zur Logik empirischer Theorien*, Berlin: Walter de Gruyter & Co., 1983.
Balzer, W. & J.D. Sneed, 'Generalized Net Structures of Empirical Theories', *Studia Logica, 36*, 1977, pp. 195-211; *Studia Logica, 37,* 1978, pp. 167-194.
Bourbaki, N., *Theory of Sets*, Paris: Herman, 1968.
Burt, R.S., 'Models of Network Structure', *Annual Review of Sociology, 6,* 1980, pp. 79-141.
Coleman, J.S., *Power and the Structure of Society*, New York: Norton, 1974.
Kuhn, T.S., *The Structure of Scientific Revolutions,* 2nd ed., Chicago: The University of Chicago Press, 1970.